DOCUMENT RESUME

ED 177 172 TM 009 487

AUTHOR Karma, Kai
TITLE The Ability to Structure Acoustic Material as a Measure of Musical Aptitude. 4. Experiences with Modifications of the Acoustic Structuring Test. Research Bulletin. No. 51.
INSTITUTION Helsinki Univ. (Finland). Inst. of Education.
PUB DATE Dec 78
NOTE 59p.; For related documents, see ED 092 440, ED 110 402, and ED 134 485; Best copy available

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Adults; Aptitude; *Aptitude Tests; Cognitive Ability; Comparative Testing; Complexity Level; Elementary Secondary Education; Factor Analysis; Foreign Countries; Item Analysis; *Listening Tests; *Music; Test Construction; Test Interpretation; Test Reliability; *Tests; Verbal Ability
IDENTIFIERS *Acoustic Structuring Test; Finland; Spatial Ability

ABSTRACT

Four new versions of an acoustic structuring test were developed, administered, and analyzed in order to produce better tests and to contribute to better understanding of the abilities measured by these tests. The tests consist of tape recordings of patterns of musical notes played on an electric organ or an acoustic guitar. Item analyses and factor analyses were used in the studies of these tests, and comparisons were made with earlier test forms. Several different item formats were used. The acoustic structuring tests were correlated with verbal and spatial ability tests, but the results of different studies appeared to conflict with each other. Seven hypotheses concerning characteristics of items that might contribute to item difficulty were studied. Items were categorized with respect to these seven characteristics, and regression analyses were conducted to find the effect of these various item characteristics. (Nine pages of music which illustrate the test items are appended.) (CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED177172

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

RESEARCH BULLETIN

Institute of Education

University of Helsinki

SF-00100 Helsinki 10

Finland

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Kai Karma
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

No. 51

December 1978

Kai Karma

THE ABILITY TO STRUCTURE ACOUSTIC MATERIAL AS A MEASURE OF MUSICAL APTITUDE
4. Experiences with Modifications of the Acoustic Structuring Test

TM 009 487

RESEARCH BULLETIN

Institute of Education, University of Helsinki
Head: Erkki A. Niskanen, Professor of Education
Fabianinkatu 28, SF-00100 Helsinki 10, Finland

Kai Karma

THE ABILITY TO STRUCTURE ACOUSTIC MATERIAL AS A MEASURE OF MUSICAL APTITUDE

4. Experiences with Modifications of the Acoustic Structuring Test

Helsinki 1978

ISBN 951-45-1559-5
ISSN 0073-179X

Contents:

page

1. Versions D and E 1
   1.1. Test construction 1
   1.2. Collecting the material 2
   1.3. Item analyses and reliabilities 3
   1.4. Relations to other variables 5
2. Versions F and G 8
   2.1. Test construction and material 8
   2.2. Item analysis and reliability 9
   2.3. Relations to other variables 10
3. Item analyses over the different versions of the acoustic structuring test 15
   3.1. What makes an item difficult? 15
   3.2. Are there subscales in the test? 22

References 29

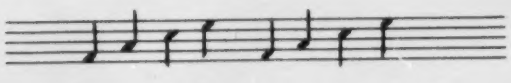
Appendix 1 31
Appendix 2 34
Appendix 3 37
Appendix 4 40
Appendix 5 41
Appendix 6 42
Appendix 7 43
Appendix 8 44
Appendix 9 45
Appendix 10 47
Appendix 11 48
Appendix 12 49
Appendix 13 50
Appendix 14 51

1. Versions D and E

1.1. Test construction

The A-, B-, and C-versions of the acoustic structuring test had an unsatisfactory reliability but gave an impression of a relatively good validity in the limits allowed by the reliability (Karma 1975). The main objective in later test construction was thus raising the reliability of the test without affecting its general character. Theoretically, the simplest way of doing this would have been making the test longer, but this has been avoided through the whole project because younger subjects get easily bored and this in turn lowers the value of the test results. Another obvious way of making a test more reliable is to encrease the number of alternatives per item. A way of doing this has been suggested earlier (Karma 1975, 17); this method was used in composing version D.

As in the earlier versions, the items of version D consist of repetitions of small figures of tones. The patterns are formed by changes in pitch, length, or intensity of the tones; only one of these vary in each item (appendix 2). The task of the subjects is to detect the amount of sub-groups within the items. For example, the correct answer to the following item would be "two" because there are two similar subgroups in it:



The number of subgroups varies from one to four which doubles the amount of alternatives when compared with the earlier versions of the test.

The tape was made by cutting and gluing from tones played with an electric organ. The timbre and the amount of simultaneous tones were varied between the items to motivate the subjects, especially the young ones. The basic rhythmic unit (marked by a crotchet (♩) is .63 seconds long (12 cm of tape in 19 cm/sec speed). As in the earlier versions, measuring of sensory capacities was avoided by making all structure-forming differences big enough to be clear to a normal person. The six example items were presented both visually and acoustically to the subjects.

When version D had been tested in practice, version E was made by rearranging the items of version D in an approximate order of difficulty and leaving one item away. These two versions are thus very similar and their results are comparable with each other.

1.2. Collecting the material

Version D of the acoustic structuring test was used in connection with the DPA Helsinki-project led by prof. Koskenniemi (DPA for Didactic Process Analysis, see, for instance, Koskenniemi - Komulainen 1977, Koskenniemi et al 1977). A relatively large amount of data is thus available about the subjects including the factor ability test battery by Heinonen (1963), Raven's Progressive Matrices (1958), the vocabulary part from "lukima"-test (Tasola 1970), information about school achievement and the like.

Six fourth-grade public school classes were tested during the academic year 1974 - 1975 the total numerus being 172. Five of the classes were situated in different parts of Espoo (a neighbour city of Helsinki with about 130 000

inhabitants) the sixth being the class of the Institute of Education, University of Helsinki. Testing was spread over a time of some months to avoid the risk of making the subjects bored. All tests were given to one class at a time; tests were given individually only in cases where a subject had not been present in the group testing situation.

Version E was given to the applicants for the Music Institute of Espoo in the spring 1975. As to the age, the applicants varied from six years olds to adults. The test was given to groups of about twenty. When the results of those young subjects who clearly had not understood the instruction were excluded, a total of 269 was left. In addition to this group test, the applicants took individually a test where they sang a tune, played something if they could and imitated some rhythms, intervals and melodies. For practical reasons, no other tests could be given.

1.3. Item analyses and reliabilities

When consulting the tables 1 and 2 one can see that the general impression is much the same. In both cases the item-total correlations are positive and most of them are on a satisfactory level. In these correlations, the total includes also the item in question; this way of calculating is simpler and was considered exact enough. "True" item-total correlations, where every item is removed in turn, would be slightly lower.

The differences between the tables may be explained on the basis of differences in the groups tested. Table 1 shows consistently lower p-values, i.e. the test has been more difficult for the unselected public school material than for the applicants for music training. Obviously, the most difficult items have been too difficult for most of the unselected subjects which has lowered the corresponding correlations. The difference between the two materials

has probably affected the reliabilities via differences in the variances. Version D which had the bigger variance gave an alpha coefficient of .78 the corresponding figure for version E being .74. For such short tests these values may be considered relatively good.

Table 1. Item-total correlations (r_{it}), p-values (p) and standard deviations (s). Version D. N=172.

item	1	2	3	4	5	6	7	8
r_{it}	.306	.418	.208	.600	.530	.504	.324	.540
p	.90	.77	.51	.78	.75	.83	.54	.77
s	.30	.42	.50	.42	.43	.37	.50	.42

item	9	10	11	12	13	14	15	16
r_{it}	.344	.381	.152	.507	.410	.479	.151	.195
p	.59	.41	.33	.72	.77	.44	.34	.65
s	.49	.49	.47	.45	.42	.50	.48	.48

item	17	18	19	20	21	22	23	24
r_{it}	.562	.054	.013	.437	.412	.412	.462	.376
p	.79	.30	.43	.35	.78	.74	.63	.33
s	.41	.46	.50	.48	.42	.44	.48	.47

item	25	26	27	28	29	30	31	32
r_{it}	.259	.311	.289	.076	.471	.212	.298	.162
p	.29	.38	.56	.17	.65	.57	.56	.40
s	.46	.49	.50	.38	.48	.50	.50	.49

item	33	34	35	36
r_{it}	.510	.080	.421	.479
p	.77	.14	.57	.39
s	.42	.34	.50	.49

Table 2. Item-total correlations (r_{it}), p-values (p) and standard deviations (s). Version E. N=269.

item	1	2	3	4	5	6	7	8
r_{it}	.205	.245	.247	.168	.259	.258	.410	.302
p	.97	.99	.96	.98	.94	.98	.97	.92
s	.18	.09	.21	.14	.23	.15	.17	.27

item	9	10	11	12	13	14	15	16
r_{it}	.398	.325	.430	.354	.387	.164	.374	.194
p	.78	.98	.61	.94	.96	.89	.92	.96
s	.41	.14	.49	.24	.21	.32	.27	.20

item	17	18	19	20	21	22	23	24
r_{it}	.268	.273	.379	.337	.234	.325	.413	.590
p	.93	.65	.74	.83	.96	.91	.77	.88
s	.25	.48	.44	.37	.20	.29	.42	.32

item	25	26	27	28	29	30	31	32
r_{it}	.381	.483	.435	.167	.382	.461	.444	.043
p	.60	.86	.78	.62	.88	.61	.88	.37
s	.49	.35	.42	.49	.33	.49	.33	.48

item	33	34	35
r_{it}	.478	.311	.384
p	.59	.62	.70
s	.49	.49	.46

1.4. Relations to other variables

From the beginning of the project the relation between musical and spatial abilities has been a point of special interest. At first a hypothesis was formed which states

that the relation between musical and spatial ability is stronger than the relation between musical aptitude and general intelligence (Karma 1973). The phenomenon of laterality and the common finding that musical ability is mainly located in the right hemisphere (Scheid - Eccles 1975, Rennels 1976, Bannatyne 1971, 212 - 241) made it seem more exact and useful to change "general intelligence" to "verbal ability"; the hypothesis is used in this form here. The data collected in connection with the entrance examinations of the Music Institute of Espoo supported this hypothesis (Karma 1975).

The sum of two subtests from Heinonen's battery having high loadings on the verbal factor (sentence completion and opposites) was used as a measure of verbal ability in connection with the version D of the acoustic structuring test. The sum of mirror test and square completion was used to operationalize the spatial factor. The correlations were as follows: spatial/musical .32 and verbal/musical .47. In the above mentioned data from the entrance examinations of the music institute the corresponding correlations were .33 and .09 which means that there has been a significant raise in the verbal/musical correlation. The change was somewhat surprising and made the validity of the new versions suspicious.

Version D was also factor-analyzed with the other ability tests made to the subjects. The four-factor solution was the clearest of the orthogonal varimax rotations available. Relatively pure verbal, spatial, and numerical factors emerged and a less clear one which was interpreted as the reasoning factor.

The structuring test gets an unexpectedly high loading on the reasoning factor. This may be thought to be a consequence of the counting operation involved; there may be a smaller possibility to come to the right solution intuitively and spontaneously in versions D and E than in

the earlier versions of the test. This, too, was considered a possible sign of low validity.

Table 3. Factor matrix. The acoustic structuring test, version D, with some ability tests. Varimax solution. N=172.

I	II	III	IV	h^2	
.148	.266	.464	.138	.327	additions
.202	.648	.103	.145	.493	perception speed
.590	.254	.236	.449	.671	opposites
.077	.639	.088	.084	.429	square completion
.707	.179	.193	.263	.639	sentence completion
.605	.306	.230	.428	.697	logical order of words
.661	.273	.132	.220	.577	word groups
.349	.127	.587	.167	.510	mathematical problems
.132	.600	.197	.217	.464	mirror test
.542	.054	.036	-.021	.299	ordering letters to words
.721	.080	.363	.174	.688	synonymes
.570	.220	.290	.331	.567	word analogies
.552	.048	.209	.504	.605	vocabulary (lukima)
.130	.403	.324	.527	.561	Raven's progressive matrices
.203	.268	.093	.474	.346	structuring test, version D
3.362	1.835	1.154	1.521	7.872	eigenvalues

I verbal II spatial III numerical IV reasoning

On the other hand, when version E was given to the applicants for the Music Institute of Espoo, correlations with playing, singing, and the imitating test developed at the institute did not differ from those attained with the earlier versions. In all, the versions D and E were considered less successful, however, and work on new ideas was begun.

Table 4. Versions A, C, and E correlated with playing, singing, and the imitating test. Data from five different entrance examinations.

	version of the acoustic structuring test and year of testing				
	A 1973	C 1974	C 1976	C 1977	E 1975
playing	.20	.24	.27	.30	.17
singing	.06	.12	.11	.10	.12
imitating test of the institute	.26	.33	.41	.15	.25

2. Versions F and G

2.1. Test construction and material

Version F is based on version C but the way of answering has been changed. In version C the subject hears a succession of sounds constructed so that it can be divided into three similar subgroups. After that follows a smaller group of sounds which may or may not be similar to the subgroups of the first part. The correct answer must be chosen from two alternatives, "yes" or "no".

In versions F and G there is no "answer" part after the first group of sounds but its end may have been changed by adding one tone to it or taking one tone away. The subject must choose from three alternatives, "one tone added", "no change" or "one tone taken away". For example, the right answer to the following example would be "one tone added":

Version F was made by cutting and gluing from a tape recorded with an electric organ in the same way as versions D and E. A pilot study was done in one fourth-grade class and some corrections were made on the basis of an item analysis; version G was made by recording the corrected result with an acoustic guitar. The instrument was changed because of practical reasons, mainly lack of time. The direct playing without cutting the tape may have caused small inexactnesses in the lengths or intensities of the tones but this was not noted to be essential. Because version F was used in a small pilot study only it is not handled exhaustively here; only version G is written in notes (appendix 3). Version G was given to three public school fourth-grade classes in Espoo in the spring 1977 the total numerus being 83. Some tests from Heinonen's factor test battery were also given to the same subjects. Tests were given to groups consisting of one class at a time; no individual testing was done.

2.2. Item analysis and reliability

In an average, the item-total correlations are on the same level as in versions D and E. Again, the most difficult items seem to be the worst, one even has a negative relation to the total score. In the case reported here the reliability expressed as the alpha coefficient is .68. By replacing the most difficult items by easier ones it should be possible to raise the reliability over .70; this is an evident next step in test development.

The reliabilities of the different versions have followed closely the number of alternatives per item in each version. Three alternatives per item seems to be the best solution constructed this far: its reliability can probably be raised to, or almost to, the level of the four-alternative versions, and the dubious counting operation can be left off. According to Grier (1975) three alternatives per item also maximize

the reliability for a fixed total amount of alternatives in the test and make the test optimally discriminative.

Table 5. Item-total correlations (r_{it}), p-values (p) and standard deviations (s). Version G. N=83.

item	1	2	3	4	5	6	7	8
r_{it}	.370	.358	.453	.471	.279	.379	.528	.252
p	.95	.92	.58	.94	.89	.62	.68	.91
s	.22	.27	.50	.25	.32	.49	.47	.29
item	9	10	11	12	13	14	15	16
r_{it}	.507	.460	.599	.412	.375	.038	.163	.359
p	.95	.90	.72	.41	.94	.59	.89	.67
s	.22	.30	.45	.49	.25	.49	.32	.47
item	17	18	19	20	21	22	23	24
r_{it}	.205	.195	.479	.339	-.113	.040	.326	.164
p	.91	.34	.87	.44	.22	.10	.71	.22
s	.29	.48	.33	.50	.41	.30	.46	.41
item	25	26	27	28	29	30	31	32
r_{it}	.365	.047	.539	.021	.235	.462	.144	.236
p	.28	.77	.75	.08	.20	.82	.27	.59
s	.45	.42	.44	.27	.40	.38	.44	.49

2.3. Relations to other variables

Spatial and verbal abilities were operationalized with the tests used already in connection with version D (sentence completion plus opposites for verbal ability and mirror test plus square completion for spatial ability). Thus the results were as much comparable as possible. The correlations were as follows: spatial/musical .06 and verbal/musical .45.

Again, this result is in a clear contrast with the earlier ones attained when applicants for music training were used as subjects. Two explanations for this seem possible: first, changes in the musicality test might have caused the differences in the correlations. Second, the subjects may have been different in the different testing situations. Neither of these explanations could be preferred because the changes in the test had been simultaneous with the changes in the subject groups. Therefore, the subjects tested with version G were also tested with the version C of the structuring test. The correlations were as follows: spatial/musical -.09 and verbal/musical .43. Table 6 summarizes the situation at this stage.

Table 6. Correlations between spatial and verbal abilities and some versions of the structuring test.

year of testing	material	correlation
1973	applicants for music training	
	version A/spatial ability	.33
	version A/verbal ability	.09
1974	applicants for music training	
	version C/spatial ability	.33
	version C/verbal ability	.09
1975	public school fourth-graders	
	version D/spatial ability	.32
	version D/verbal ability	.47
1977	public school fourth-graders	
	version G/spatial ability	.06
	version G/verbal ability	.45
	version C/spatial ability	-.09
	version C/verbal ability	.43

The table suggests strongly that changes in the subject groups have caused the changes in the correlations. It seems obvious that the original hypothesis stating that musical aptitude correlates more with spatial than verbal ability was too simple. A more complicated hypothesis was then formed: it is possible to structure acoustic material in (at least) two different ways: the "verbal" and the "spatial". What the empirical correlations will be like depends on the composition of the subject group tested. Certain background variables explain which "listening strategy" an individual chooses.

Which might these background factors be, then? The correlations reported above give one possible answer to this. When the subjects have been applicants to the music institute, there has been a closer connection between spatial and musical than verbal and musical abilities. In a nonselected sample the relations seem to be the other way round, however. The big amount of applicants has forced the music institute in question to be very selective; this in turn means that most of the applicants already have former music training. It seems reasonable to suppose that music training in its usual form increases the correlation between sound organizing and spatial imagery. Quite a lot of the terminology used is directly spatial in nature;for instance "descending melody", "high tones", and "intervals between tones" are such expressions. The whole system of notation is a spatial analogy of a thing not originally spatial. It may also be that the mere training of the "right hemisphere abilities" (mostly non-verbal) increases the capacity to use this kind of thinking in different situations. It is often argued that the usual school ignores this sice to a great extent (For instance Rennels 1976).

Sex is known to be a factor which is related to spatial and verbal abilities. Girls are usually better in verbal tasks and, on the other hand, boys tend to be better in spatial problems.

This is also true here. In the 1975 and 1977 materials which are the most complete the figures are as follows:

Table 7. Verbal and spatial scores for girls and boys in the 1975 and 1977 materials.

		verbal score (=opposites + sentence completion)		spatial score (=mirror-test + square completion)	
1975	girls	M=37.1	s=12.3	M=20.1	s=9.5
	boys	M=32.9	s=15.3	M=21.8	s=9.1
1977	girls	M=48.4	s=12.2	M=27.0	s=8.2
	boys	M=39.6	s=11.9	M=29.2	s=8.4

It is not quite clear why the 1977 subjects are better in an average than the 1975 group. The -75 group was tested in the beginning and in the middle of the fourth school year, the -77 group was tested at the end of the academic year, but this explains the difference only partially. The relations within each group are the main objects of interest, however.

Supposing that the two different ways of structuring acoustic material exist and knowing that girls are more verbal and boys more spatial, one is easily led to the thought that girls use more their verbal ability and boys their spatial ability in solving acoustic structuring problems. This hypothesis is not supported by this material, however. When sex is used as the basis of classification, the correlations are as follows in the -75 and -77 materials:

Table 8. Spatial/musical and verbal/musical correlations in the 1975 and 1977 materials classified according to sex.

		$r_{spat/mus}$	$r_{verb/mus}$
1975 [1]	girls	.33	.55
	boys	.31	.41
1977 [2]	girls	-.02	.37
	boys	.11	.40

1) musical score = version D of the acoustic structuring test
2) musical score = version C + version G

The persons tested were also classified as "spatial" and "verbal" types according to which of the corresponding scores had the higher Z-value. The result of this analysis was similar to the one above: it gives the impression that spatial/verbal dominance does not affect ones listening strategy. The difference between the applicants for music training on one hand and unselected public school pupils on the other seems to be the only explanation available here.

Table 9. Spatial/musical and verbal/musical correlations in the 1975 and 1977 materials classified according to spatial/verbal dominance.

		$r_{spat/mus}$	$r_{verb/mus}$
1975	verbal group	.39	.47
	spatial group	.44	.52
1977	verbal group	.06	.41
	spatial group	.25	.40

3. Item analyses over the different versions of the acoustic structuring test.

In addition to such standard operations as calculating the item-total correlations, the different versions of the acoustic structuring test were also analyzed in other ways. The analyses discussed in these two chapters are attempts to get information about the thought processes needed in solving the items. Because five different versions of the test are analyzed and the empirical results have been obtained from five different subject groups making together a total numerus of 1098 the results may be considered generalizable to some degree.

Tracking the subjects' thought processes could also serve as one approach in estimating the validity of the test. The test is intended to measure the ability to structure acoustic material. Showing, for instance, that the items can be answered correctly by using mere sensory discrimination or rote memory would be showing that there are serious faults in the test. On the other hand, showing that abilities that are connected with structures are used would give support to the validity of the test.

The subjects' thought processes could, of course, be investigated by asking what they were thinking during the testing. This has not been done this far although it might be an approach worth trying. In most testing situations the practical limits have made individual interviews impossible and, moreover, this strategy would give information about conscious thinking, only.

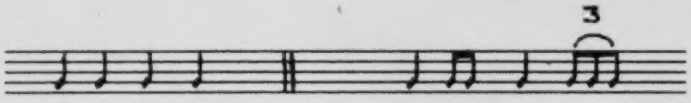
3.1. What makes an item difficult?

The statistical procedure chosen here was to study which properties of the items covary with the levels of difficulty

of the items. Items of several difficulty levels make finer differentations between the subjects possible; making these differentations is measuring. From the point of view of measuring, the critical variables are the ones which cause the different difficulty levels; these should correlate with the difficulty. The concrete statistical problem is thus explaining the difficulty with other properties of the items.

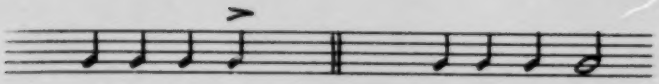
Versions A, C, D, E, and G were analyzed in this way; versions B and F were omitted because of their use in pilot studies and with small subject groups, only. Regression analysis was used as the statistical method, the criterion was the empirically obtained difficulty, i.e. the q-values of the items. The predictors were as follows:

1) The length of one subgroup, i.e. how many notes there are in it. Counting these rose a question of the possible mechanical nature of this measure. Perhaps the human mind interpretes subdivisions of the basic rhythm as different kinds of tones instead of different amounts of them. If so, both of the following groups would consist of four elements, the latter being a variant of the former, basic one:



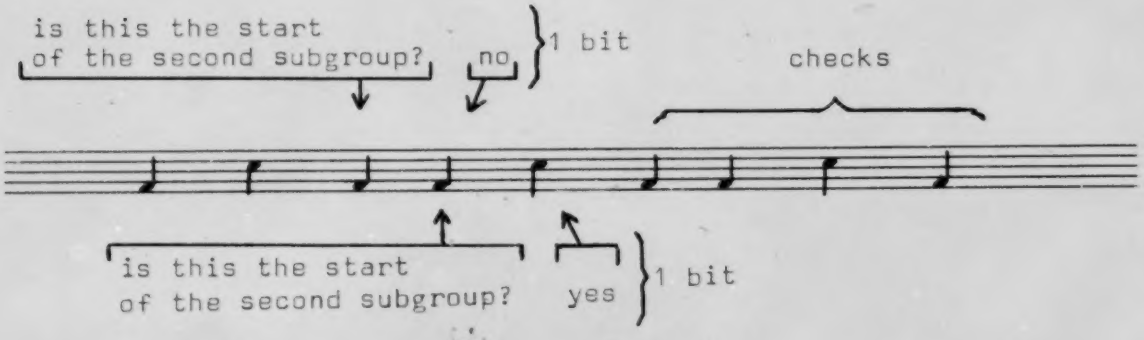
Both ways of counting the amounts of tones were tried but the difference in the correlations was negligible. The total number of tones is used in the computations here; the only exceptions are the triplets in versions A and C, which were interpreted as single elements.

2) Misleading accent. An accent in the end of one subgroup easily masks the beginning of another subgroup. In addition to "real" accents where the intensity is increased, also longer notes were interpreted as accents. For instance, the following groups are both considered to have misleading accents:



There was room for some subjective judgement in this variable, too. Cases in which there is an accent both in the beginning and in the end of a subgroup, as well as those where the accent is in the middle, are not clear. The difference between the different interpretations was very small in the correlations, however; only the cases where the accent is in the end of a subgroup were considered misleading accents in the final calculations.

3) The minimum amount of "bits" of information needed to make sure how the group of sounds is constructed. In forming this variable, some assumptions about the thought processes needed were made. It seems inevitable that one must - consciously or inconsciously - wait for a sound similar to the one the item began with. This is the first possible starting point of a new subgroup. In other terms, this is the first point of uncertainty where the subject must in a way or another ask himself: "Is this the beginning of the second subgroup?". If the following tone is similar to the second tone of the whole item, the answer is "yes" and one "bit" of information - in the a little imprecise meaning the term is used here - is used for solving the item. The rest of the sounds may be considered "checks" which support the decision. If, however, the expectation is not filled, the answer is "no". Again, one bit is used but the problem remains unsolved. The same procedure must be gone through at the following point of uncertainty and so on. The principle may be clarified with the following example of a "two bit" item:



The concept of bit as it is used here may be said to measure the flexibility of structure formation, the ability to change a conceived structure when needed. The possible importance of this variable has been noted earlier (Karma 1973, 17).

4) "No" item. When the items of a test are dicothomous, i.e. when there are two alternatives to choose from, the wrong alternatives look right more often than vice versa. This has also been true with the dichotomous versions of the acoustic structuring test (Karma 1975, 7).

5) Pitch item. The structure-forming variables, pitch, intensity, and length could not be entered into the regression analyses together because they would have formed an ipsative group. Pitch items have been noted to be easier than length and intensity items; thus a variable "pitch vs not pitch item" was formed for the analyses.

6) Amount of different tones. In length and intensity items there are only two kinds of tones in each item: longer and shorter, louder and less loud. In pitch items the amount of tones may be bigger, however. Because this difference might explain the possible weight of pitch items as predictors, this variable was taken into the analyses.

7) Amount of subgroups. This could, of course, be used only in connection with the versions in which it varies, i.e. versions D and E.

There are some properties in the data which decrease their suitability for regression analysis. First, the numbers of statistical units which are equal to the numbers of items in each case, are relatively small (31 - 40). Second, there are some high correlations between the predictors. For instance in versions D and E the number of subgroups and the number of tones in a subgroup correlate negatively (-.71). This is a natural consequence of the fact that an item with one subgroup cannot be very short without being

self evident. Third, there are some cases where there is very little variance, especially in the variables misleading accent and amount of different tones.

For these reasons, individual analyses and significances within them are not used as usual. The different materials are used to cross-validate each other, instead. Only tendencies which show some stability over the different materials are considered important. The weights (beta coefficients) and total percentages predicted in each case (R^2) were obtained from analyses where all the predictors were included. The unique parts (U) in the prediction have been counted as the differences between the total R^2 and the R^2 without each variable in turn. In other words, they indicate how much the prediction is better when the variable in question is added to the whole. The results have been collected into table 10.

Subgroup length does not seem to have any systematic connection to item difficulty. The unique parts of this variable in the prediction are small and the signs of the weigths vary. One must remember, however, that the lengths are within the limits of the memory span of a normal person (7 - 8 units). This variable might become important in a test with longer items.

Mileading accents add very little to the total but, on the other hand, this variable behaves very similarly in the different analyses. Remembering that its variance was quite small in this material one might conclude that it tends to make structuring a little more difficult.

Amount of bits its clearly the best predictor in these analyses. Only in version A there is one that has a little higher predictive value, namely the variable pitch vs length or intensity.

Table 10. Summary of regression analyses. Item difficulty predicted with some other properties of the items.

predictors	version A		version C		version D		version E		version G	
	β	U	β	U	β	U	β	U	β	U
length of a subgroup	-.11	.009	.16	.018	.12	.004	-.26	.019	.14	.014
misleading accent	.11	.009	.16	.018	.12	.012	.13	.015	.27	.041
amount of bits	.55	.213	.55	.274	.53	.141	.69	.25	.49	.176
"no" item	.31	.082	.36	.103	---	---	---	---	---	---
pitch item	-.61	.227	.03	.000	-.31	.066	-.11	.008	-.11	.006
amount of different tones	.24	.032	-.07	.002	-.03	.001	-.12	.011	-.11	.008
amount of subgroups	---	---	---	---	-.12	.006	.06	.002	---	---
R^2	.47		.46		.58		.36		.56	

The difference between "yes" and "no" answers could be studied with two versions of the test only. The values are so close to each other, however, that it seems plausible to conclude that if other variables are constant "no" items tend to be more difficult than the ones to which "yes" is the correct answer.

The part the structure-forming variables have in the total predictions vary to such an extent that any conclusions seem hasty. There seems to be something exceptional in the version A in this respect. This is somewhat surprising because versions A and C are relatively similar. In an average, pitch items tend to be easier than length or intensity items.

The last two variables, amount of different tones and amount of subgroups, are useless as predictors. The shares they have in the total explanations are small and their signs vary.

The predictors used here are of two different kinds. First, those in which the absolute qualities of the items vary: their lengths (subgroup length and amount of subgroups), and the qualities of the sounds used (amount of different tones and pitch vs length or intensity). Second, those predictors in which the variance is mostly in the relations of the tones (misleading accent, amount of bits, "no" item). The latter group has clearly a bigger and more systematic effect on the scores of the test than the former. On the basis of this result it can be said that the tests measure relatively well what they are intended to measure: the structure-forming ability. The only exception is the first version, A, where pitch as the structure-forming variable has an unexpectedly high effect. The important role of the concept "bit" suggests that flexibility of structure forming is measured to a great extent.

3.2. Are there subscales in the test?

At an early stage of the project it became evident that theoretical development should be the primary aim before practical test development. Test constructing without a sound theoretical basis tends to become a trial and error procedure; even if the test is successful in some practical use it may be difficult to know why. In the present work it was thought that a homogenous instrument gives clearer results from the point of view of research although it may not be the best alternative for practical predictions. When the theoretical basis has been given a reasonable degree of clarity it should be much easier to compose tests for practical purposes, too. Therefore, the versions of the acoustic structuring test have been homogenized to get an effective instrument for measuring one interesting theoretical construct.

This being the case it is not probable that there would be any clear subscales in the test. When the items within one version are correlated the correlations are relatively low and their variance is small. In a situation like this the results of factor analyses tend to be unstable and difficult to interpret or even nonsense because of the big role of error variance. However, as with the regression analyses, there is a possibility to cross-validate the results of one analysis with the results of the others in this material. Thus, factor analyses are made but serious attention is paid only to tendencies which can be seen in more than one analysis. Both unrotated principal axes solutions and rotations from them were considered. The rotations did not seem to make the matrices easier to interpret in this case. Thus, principal axes solution will be discussed here paying special attention to the second factor in each solution where the common variance has been removed and properties lower in the hierarchy can be seen. Because of the nature of the principal axis method the loadings on the first factor are essentially similar to the item-total correlations and are not discussed here.

In version A there were some items which were so similar to each other that they formed a group which tended to dominate the whole factor structure. To avoid this distortion, items 10, 13, and 21 were remcved.

It was noted already in connection with the regression analyses that the expected answer in versions A and C, "yes" or "no", affected the difficulty of the item. Could it also have an impact on the factor structure? Both ends of the second factors have been collected into table 11 indicating the correct answer to each item.

It is obvious that the similarity or dissimilarity between the "answer part" of an item and the substructures in the beginning affect the factorial structure of the test. Could this external feature be explained with some internal, psychological rules or principles? It seems that it can be done.

The basic assumption needed is that two different strategies are used in solving the problems. First, the "anticipating strategy": the subject structures the first group of sounds and waits for the "answer part" asking himself in a way or another: "Will it be similar to the structure I conceived in the first part?". Second, the "reconstructing strategy": the first part leaves the subject in an uncertain state and he must ask himself after hearing the second part: "Was it in the first part?".

This assumption explains both the formation of the factors and the fact that "no" items are more difficult. If the anticipation strategy is used, there will be no essential difference between the "yes" and "no" items. If,on the other hand, part of the subjects use the reconstructing strategy, the difference is essential. In "yes" items this strategy is likely to lead to the right solution: the subject recalls the structure, answers "yes" and gets a point. In the "no"

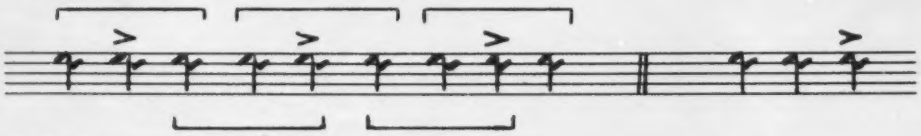
Table 11. 10 strongest positive and negative loadings on the second factors of unrotated principal axis solutions. Versions A and C.

Version A			Version C [1]		
item	loading	correct answer	item	loading	correct answer
22	.357	no	19	.331	yes
38	.353	no	6	.330	yes
11	.308	no	8	.319	yes
31	.285	no	23	.285	yes
33	.282	no	21	.284	yes
30	.280	no	31	.239	yes
20	.269	no	2	.230	yes
12	.216	no	28	.179	yes
7	.140	no	1	.162	yes
40	.065	yes	14	.139	yes
1	-.164	yes	22	-.092	no
8	-.170	yes	25	-.092	no
2	-.217	yes	15	-.150	no
25	-.229	yes	27	-.165	no
3	-.234	no	26	-.169	no
35	-.238	yes	30	-.175	no
27	-.240	yes	17	-.241	no
24	-.243	no	12	-.241	no
23	-.243	yes	3	-.255	no
6	-.343	yes	11	-.261	no

1) Version C is shown in notes in the appendix of Karma 1975

items there are many possibilities to make a mistake with this strategy. The structure of the "answer part" may exist in the first part but it is not the correct one, one third of the whole. A "shift of scale" has happened: the group of sounds has not been divided into three similar parts but in a manner which produces the wanted structure. For instance, in version A, item 22, this could happen as follows:

"correct scale":



"shifted scale":

In a case like this the subject may answer "yes" and thus be wrong. It is obvious that this produces more right answers to "yes" items, i.e. makes them easier.

In versions D and E there is no "answer" part in an item. The number of subgroups varies, instead. In table 12 ten strongest positive and negative loadings on the second factors of principal axes solutions are shown with the corresponding amounts of subgroups and numbers of tones in one subgroup.

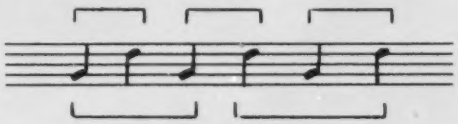
Although the results are not as clear as with versions A and C, two tendencies can be seen. First, items with two tones in a subgroup and items with three tones in a subgroup tend to form groups of their own. Second, items with one subgroup only tend to gather together. In version D the one subgroup items tend to load on the negative end of the third factor of the principal axis solution. In version E the "two tones in a subgroup"-end cannot be clearly identified although there are some traces of it on the positive end of the third factor. In spite of these irregularities these tendencies are interesting enough to be worth interpreting.

Table 12. 10 strongest positive and negative loadings on the second factors of unrotated principal axis solutions. Versions D and E.

Version D				Version E			
item	loading	amount of sub-groups in item	tones in 1 subgroup	item	loading	amount of sub-groups in item	tones in 1 subgroup
15	.442	2	3	5	.317	1	4
13	.442	3	3	30	.275	1	6
18	.427	2	3	27	.249	1	7
12	.381	2	3	26	.227	1	6
23	.375	2	3	24	.225	1	6
35	.312	3	5	35	.220	1	6
10	.282	3	3	29	.148	2	5
26	.263	3	3	23	.106	3	5
29	.251	2	4	9	.103	3	3
36	.203	2	5	14	.079	3	2
20	-.054	1	6	7	-.227	1	4
8	-.082	1	4	10	-.251	2	3
3	-.091	3	2	22	-.270	3	3
1	-.138	2	2	34	-.293	2	3
33	-.172	3	4	2	-.311	3	2
17	-.225	4	2	33	-.312	3	3
9	-.276	4	2	25	-.342	2	3
6	-.289	3	2	28	-.365	3	3
2	-.319	3	2	32	-.366	2	3
21	-.340	4	2	13	-.440	3	3

One more assumption must be made to be able to explain the "two vs three notes in a subgroup"-factor using the same kind of reasoning as with versions A and C. One must suppose that in using the "reconstructing strategy" retroversions (the original backwards) and inversions (the figures are "mirror images" of the original) are confounded with the original patterns. This assumption makes the difference between two and three notes in a subgroup understandable. Especially the items with three subgroups, each consisting of two notes are good examples of this possibility. By cutting such an item into two halves one gets two subgroups which are inversions of each other:

correct scale: three subgroups



"shifted scale": two subgroups

If there are three tones in a subgroup, inversions or retroversions cannot be formed with a "shift of scale".

Obviously, there is very little empirical evidence to support this explanation whic must be considered a hypothesis only. To get at least some evidence, the errors made in version D in items 2, 6, and 16 which are of this type, were counted. If the hypothesis holds, the most common wrong answer should be "two". The errors were distributed as follows:

	"one"	"two"	"four"
item 2	8	15	10
item 6	2	16	7
item 16	11	25	19
sum	21	56	36

The difference between the sum frequencies 56 and 36 is significant at the .05 level. The corresponding items in version E were so easy to the applicants for the music

instititute that the sum frequences were only 16, 19, and 11 out of the total of 269 tests taken. These figures form a very weak basis for any conclusions.

Although it feels subjectively understandable that the items with one subgroup may be different from the rest, it is difficult to give any general rule to explain this. The lack of possibility to check the conceived structure seems to make the situation different. Also the factor structure of version G seems to resist any clear interpretations. It seems wiser not to handle the results before additional material has been collected.

References:

Bannatyne, A. 1971. Language, Reading and Reading Disabilities, Psychology, Neuropsychology, Diagnosis and Remediation. Springfield: Charles C Thomas. Publisher.

Grier, J.B. 1975. The Number of Alternatives for Optimum Test Reliability. Journal of Educational Measurement, Vol. 12 no 2.

Heinonen, V. 1963. Älykkyystestejä lahjakkuuden eräiden alkeistekijöiden tutkimista varten. Jyväskylä: Kasvatustieteiden tutkimuskeskus.

Karma, K. 1973. The Ability to Structure Acoustic Material as a Measure of Musical Aptitude. 1. Background Theory and Pilot Studies. Research Bulletin, Institute of Education, University of Helsinki.

Karma, K. 1975. The Ability to Structure Acoustic Material as a Measure of Musical Aptitude. 2. Test Construction and Results. Research Bulletin, Institute of Education, University of Helsinki.

Koskenniemi, M. - Komulainen, E. - Kansanen, P. - Karma, K. - Martikainen, M. - Holopainen, P. - Uusikylä, K. 1977. Opetustapahtuman rakenne ja opetuksen kehys- ja ympäristömuuttujat. Didaktisen prosessin tutkimusprojektin (DPA Helsinki) yleisraportti. Tutkimuksia, Helsingin yliopiston kasvatustieteen laitos.

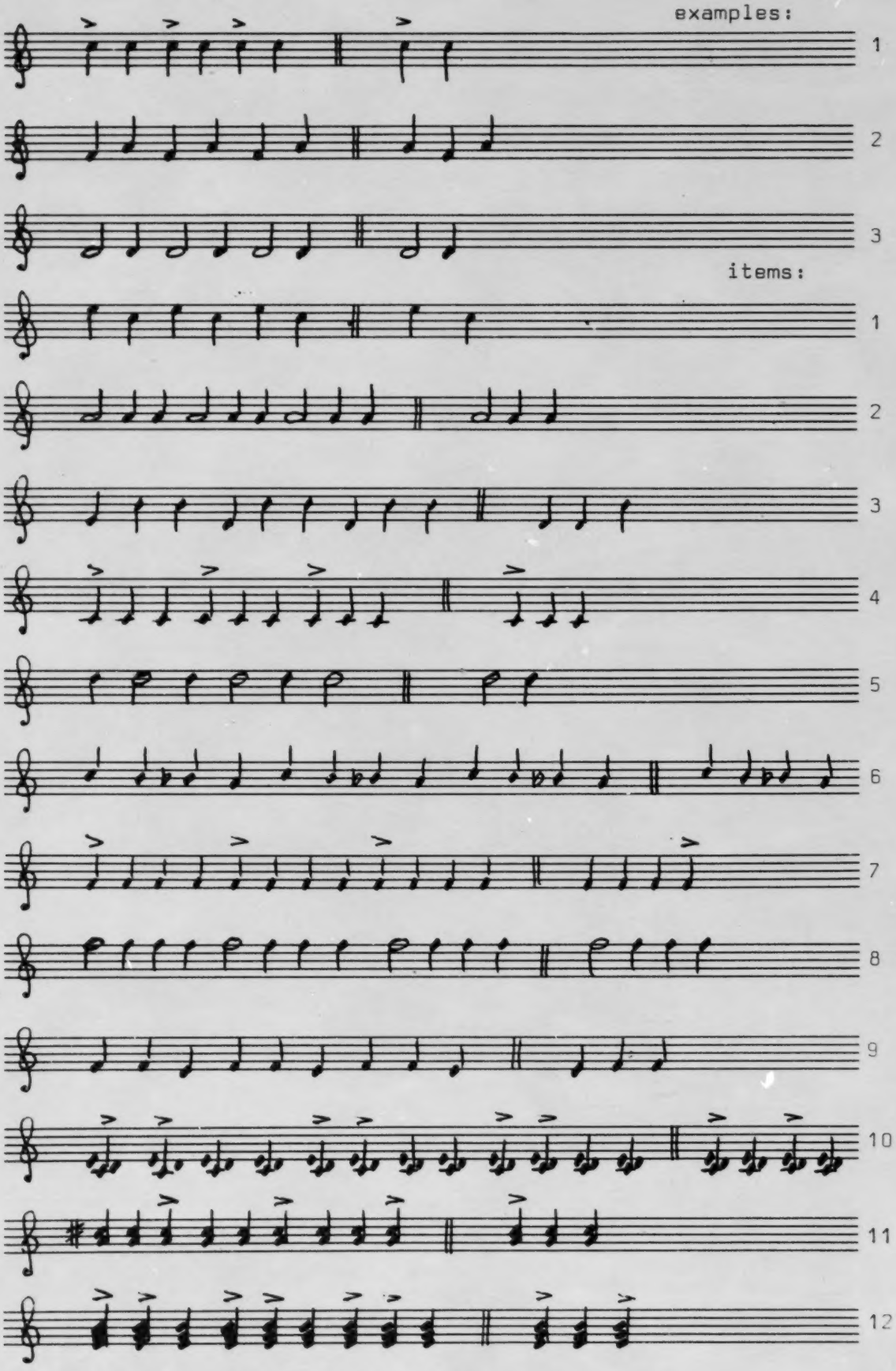
Koskenniemi, M. - Komulainen, E. 1977. Finnish Investigations into the Instructional Process (OPA Helsinki). Paper presented at 7:th World Congress of the International Association for the Advancement of Educational Research (IAAER), Gent, Belgium.

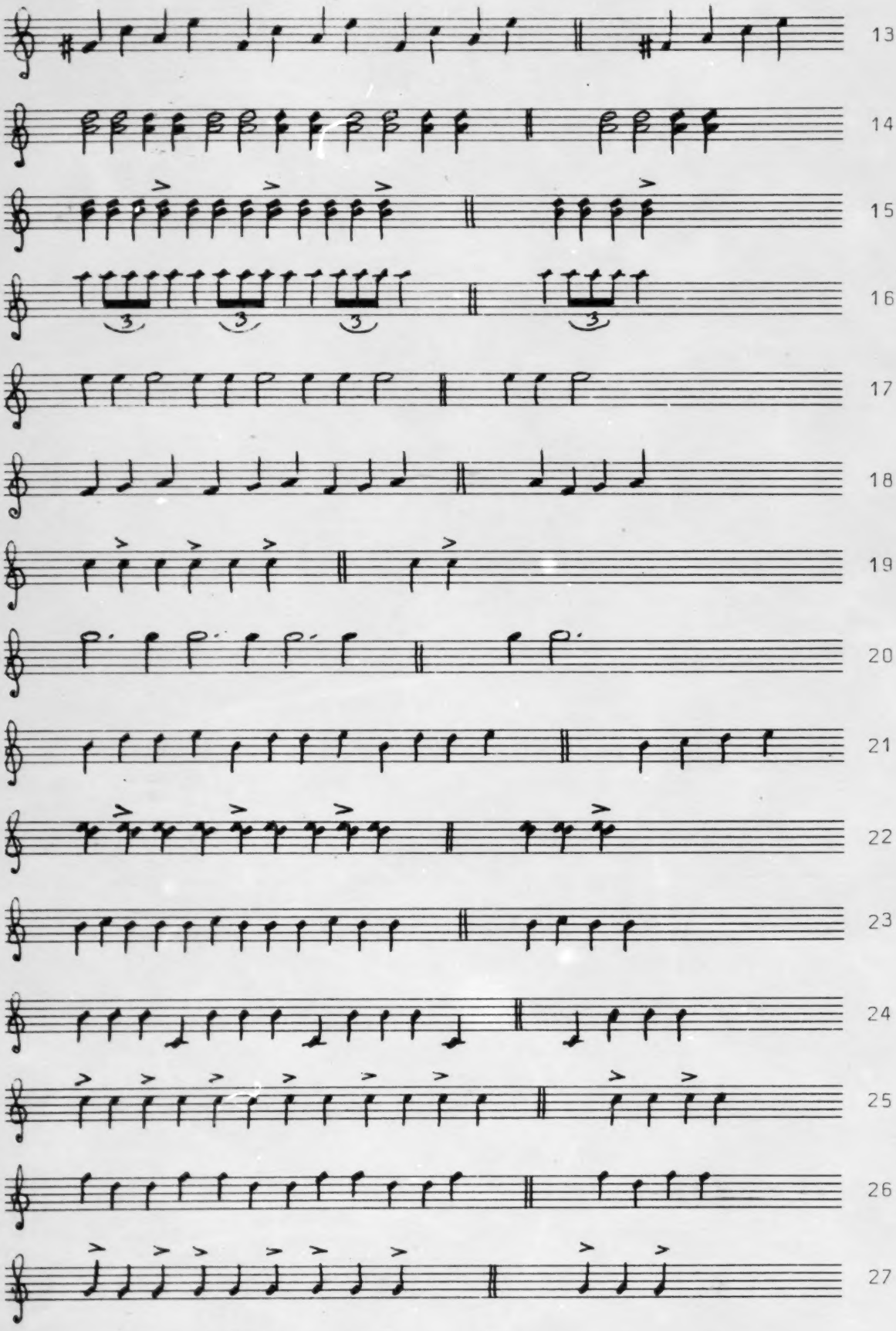
Raven, J.C. 1958. Standard Progressive Matrices. Sets A, B, C, D, and E. London: H. K. Lewis.

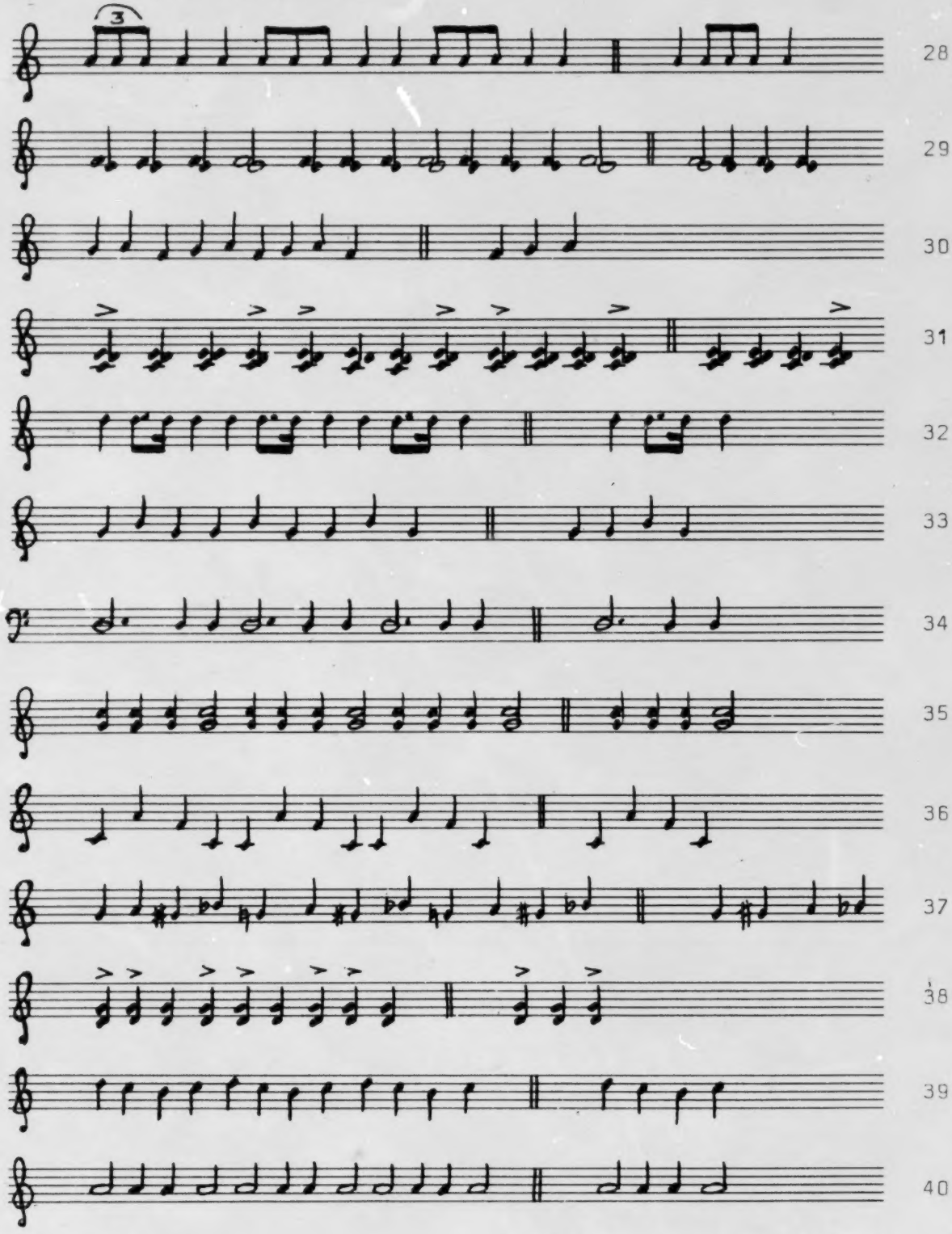
Rennels, M.R. 1976. Cerebral Symmetry: An Urgent Concern for Education. Phi, Delta, Kappan, Vol. 57 no 7.

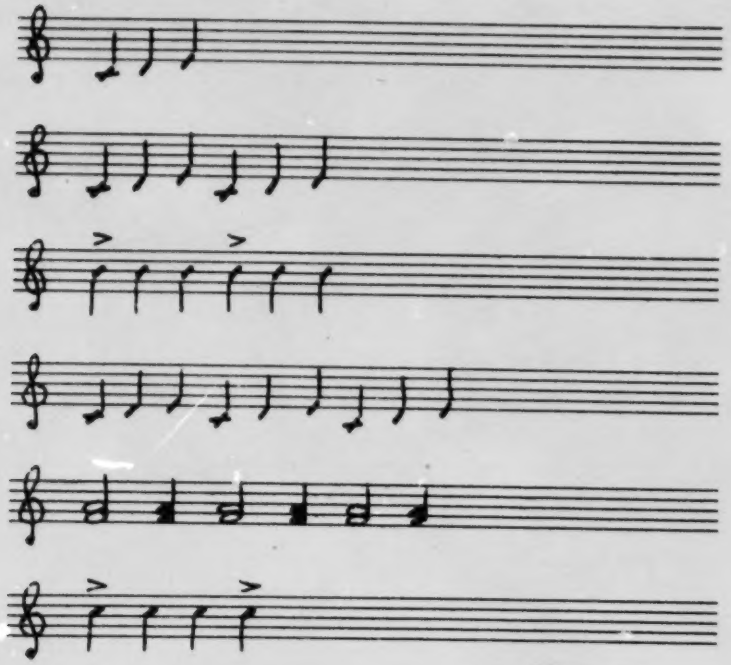
Scheid, P. - Eccles, J. 1975. Music and Speech: Artistic Functions of the Human Brain. Psychology of Music, Vol. 3 no 1.

Tasola, O. 1970. Koesarja Lukima III ja Lukima IV. Lukemisen, kirjoituksen ja matematiikan koulukokeita peruskoulun III ja IV luokalle. Opettajan ohjekirja. Jyväskylä: Kasvatustieteiden tutkimuslaitos.

Appendix 1. The acoustic structuring test, version A.

Appendix 1 continued

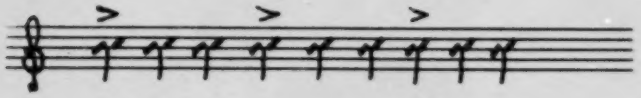
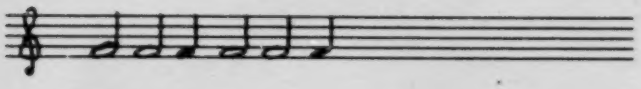
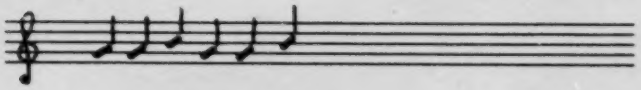
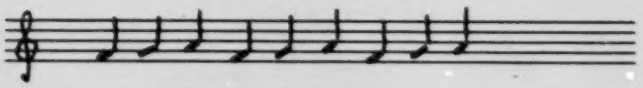
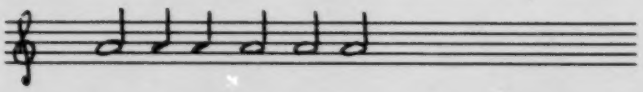
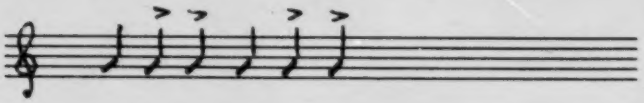
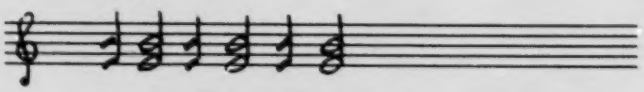
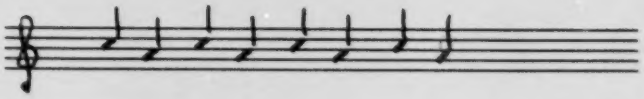
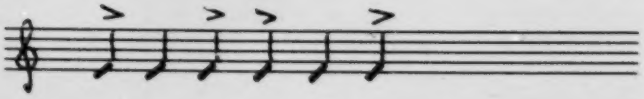
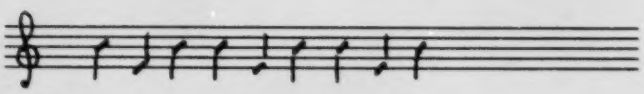
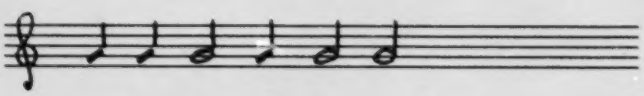
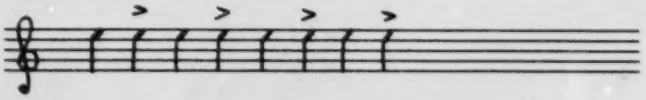
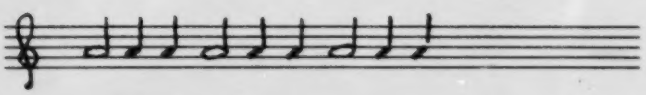
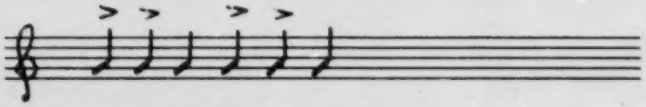
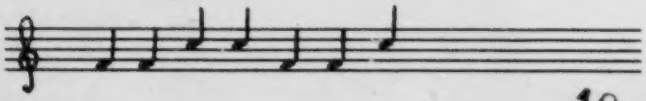
item

Appendix 1 continued

item

Appendix 2. The acoustic structuring test, versions D and E.

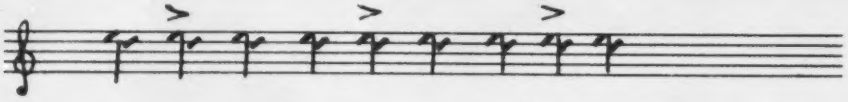
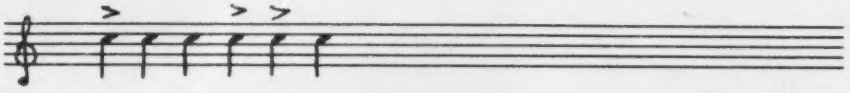
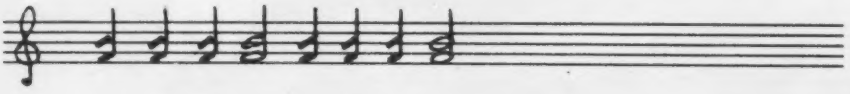
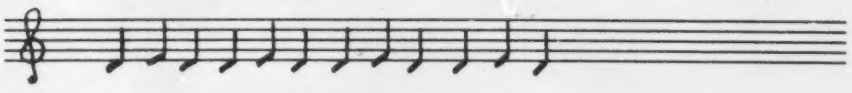
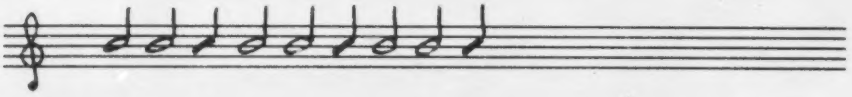
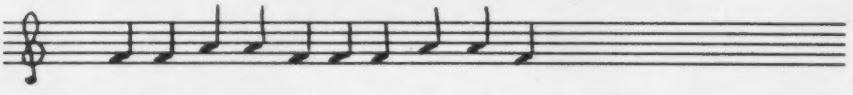
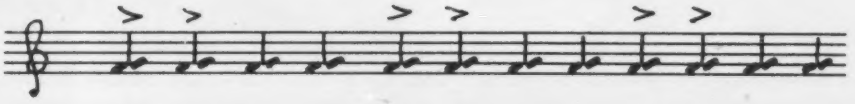
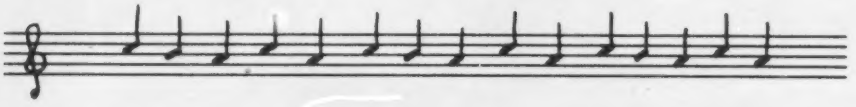
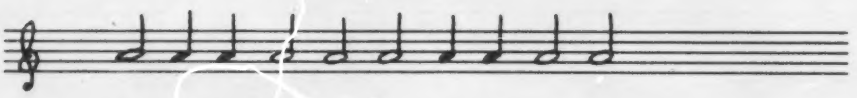
examples:

1
2
3
4
5
6

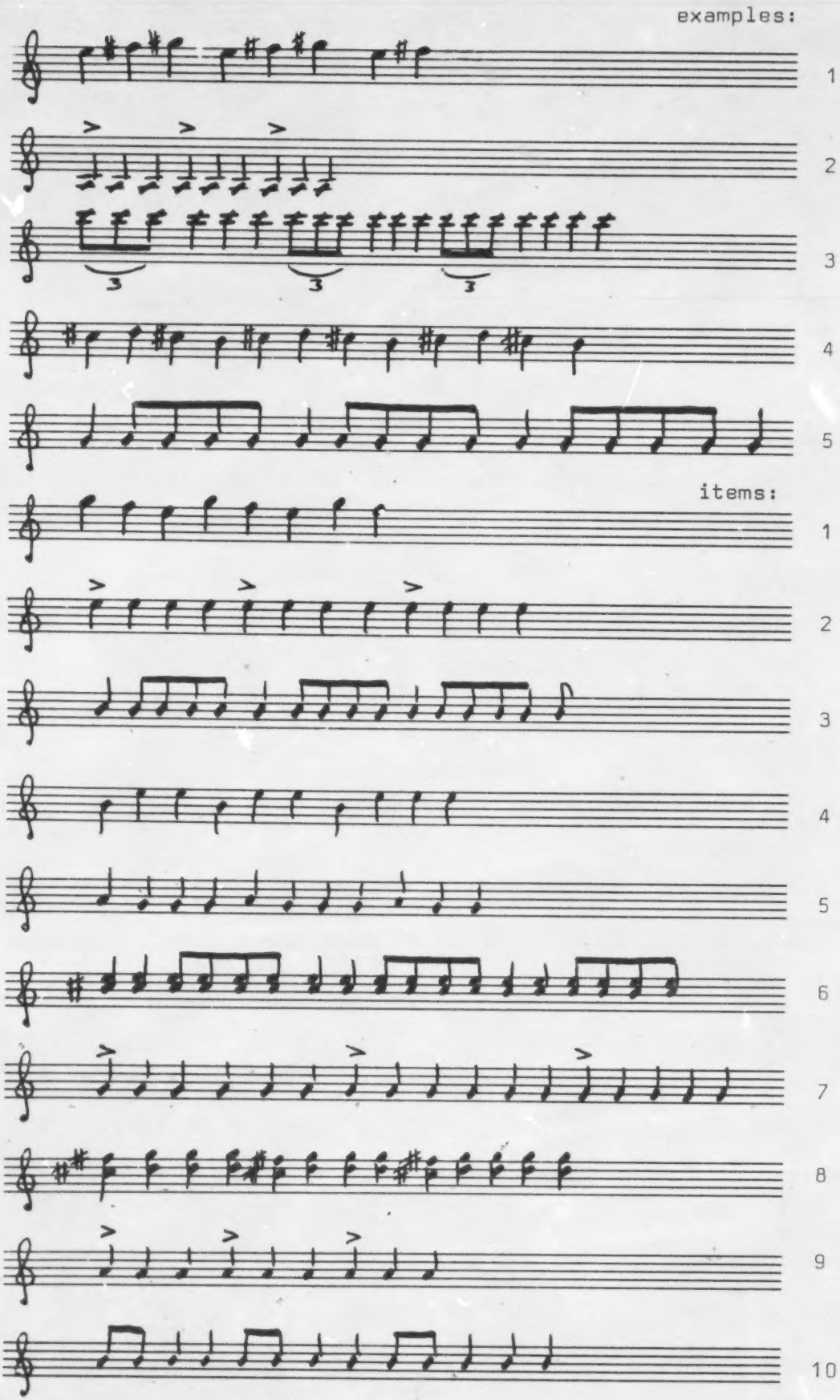
item no	version D	version E
1	1	1
2	2	3
3	3	20
4	4	4
5	5	5
6	6	2
7	7	10
8	8	7
9	9	11

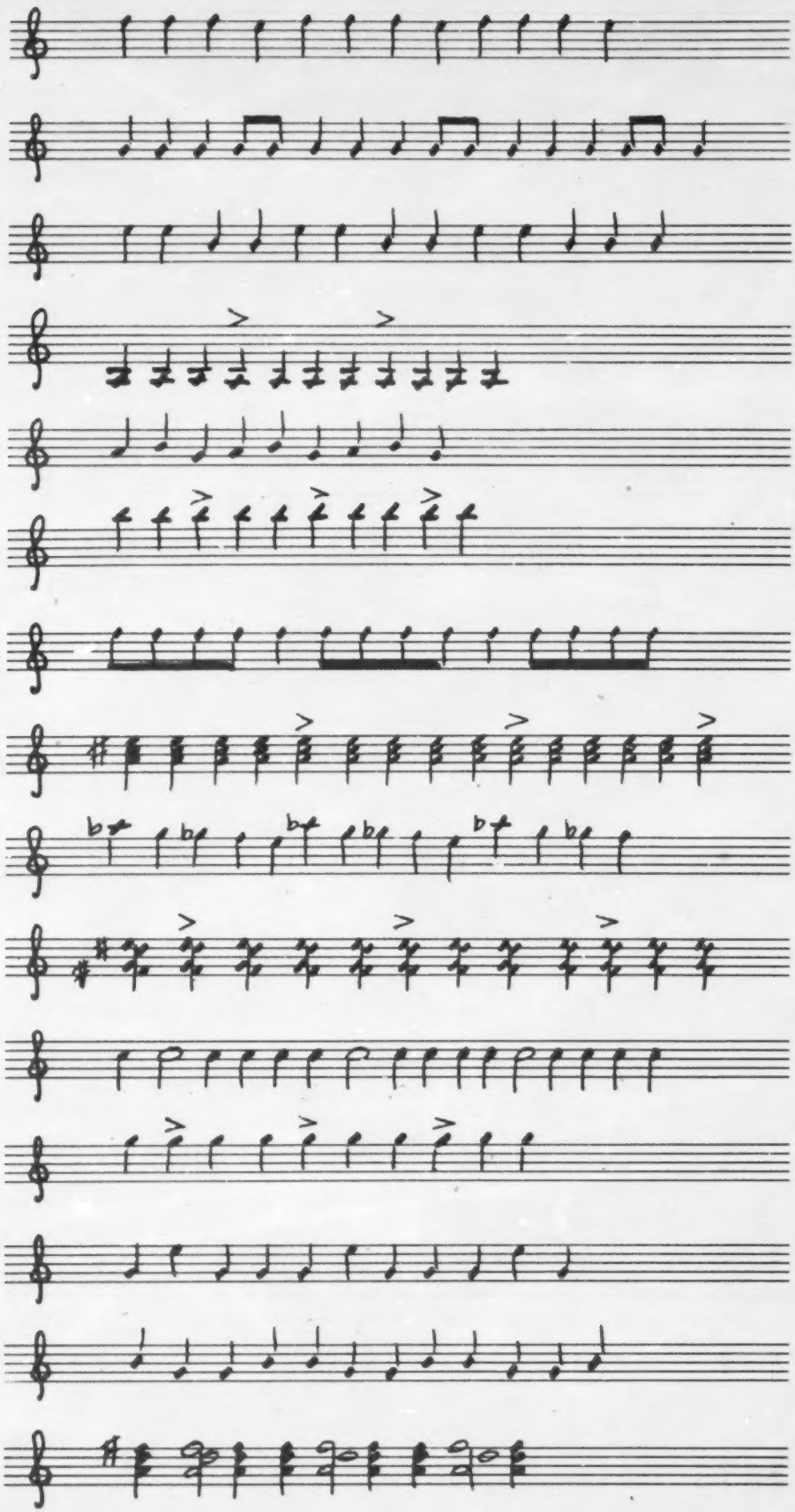
Appendix 2 continued

	item no version D	item no version E
	10	21
	11	34
	12	12
	13	13
	14	24
	15	31
	16	14
	17	6
	18	25
	19	33
	20	26
	21	16
	22	9
	23	15
	24	27

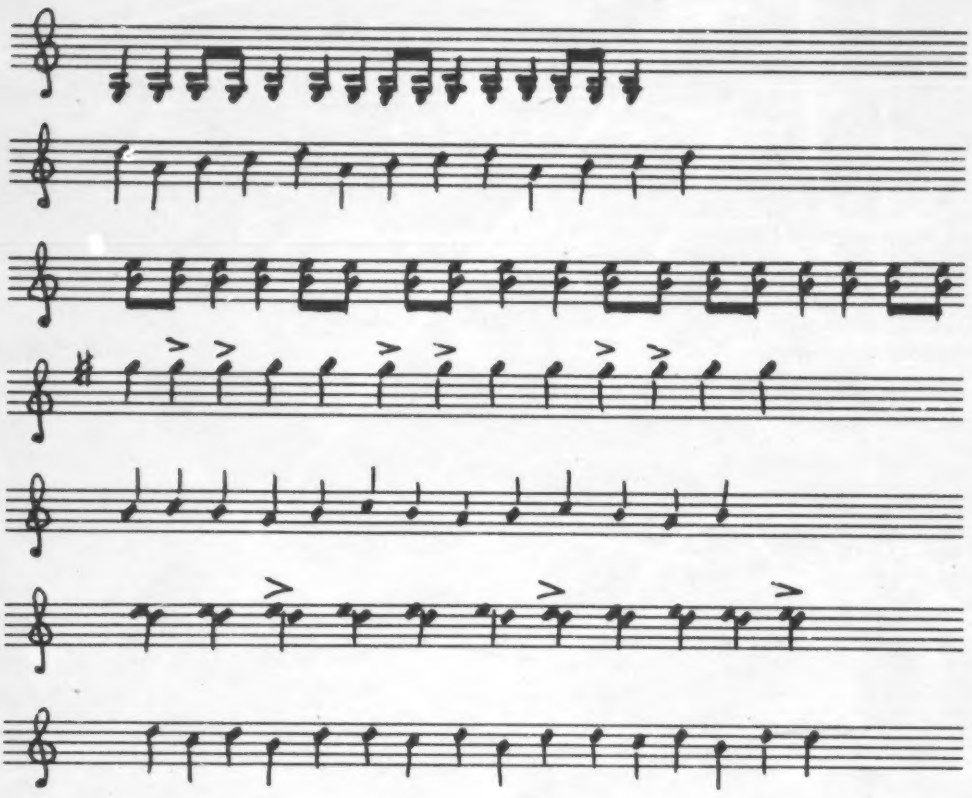
Appendix 2 continued

item no	version D	version E
	25	32
	26	28
	27	22
	28	35
	29	17
	30	18
	31	19
	32	--
	33	8
	34	30
	35	23
	36	29

Appendix 3. The acoustic structuring test, version G.

Appendix 3 continued

item

Appendix 3 continued

item



Appendix 4. Primary matrix for explaining the difficulty of the items. Version A. q-values obtained from 286 applicants for music training.

1	2	3	4	5	6	7	8
1	01	2	0	1	0	1	2
2	09	3	0	1	0	0	2
3	03	3	0	1	1	1	2
4	10	3	0	1	0	0	2
5	33	2	1	1	1	0	2
6	02	4	0	1	0	1	4
7	29	4	0	1	1	0	2
8	22	4	0	1	0	0	3
9	07	3	0	1	1	1	2
10	15	4	0	1	1	0	2
11	49	3	1	1	1	0	2
12	20	3	0	1	1	0	2
13	15	4	0	1	1	1	4
14	37	4	0	1	0	0	2
15	22	4	1	1	0	0	2
16	15	3	0	2	0	0	2
17	12	3	1	1	0	0	2
18	06	3	0	1	1	1	3
19	18	2	1	1	0	0	2
20	21	2	0	1	1	0	2
21	01	4	0	1	1	1	3
22	58	3	0	2	1	0	2
23	21	4	0	3	0	1	2
24	05	4	0	1	1	1	2
25	16	4	0	2	0	0	2
26	13	4	0	2	1	1	2
27	37	3	0	2	0	0	2
28	14	3	0	1	1	0	2
29	23	4	1	1	1	0	2
30	30	3	0	1	1	1	3
31	50	4	0	2	1	0	2
32	25	4	0	2	0	0	3
33	44	3	0	2	0	0	2
34	37	3	0	1	0	0	2
35	11	4	1	1	0	0	2
36	38	4	0	2	0	1	3
37	13	4	0	1	1	1	4
38	20	3	0	1	1	0	2
39	07	4	0	1	0	1	3
40	22	4	0	2	0	0	2

1 = number of item
2 = q-value
3 = amount of tones in one subgroup
4 = misleading accent (1 = yes, 2 = no)
5 = amount of bits (see text)
6 = "no"-item = 1, "yes"-item = 0
7 = structure-forming variable (1 = pitch, 0 = length or intensity)
8 = amount of different tones

Appendix 5. Primary matrix for explaining the difficulty of the items. Version C. q-values obtained from 309 applicants for music training.

1	2	3	4	5	6	7	8
1	01	2	0	1	0	1	2
2	05	3	0	1	0	0	2
3	01	3	0	1	1	1	2
4	05	3	0	1	0	0	2
5	11	2	0	1	1	0	2
6	03	4	0	1	0	1	4
7	07	4	0	1	1	0	2
8	16	4	0	1	0	0	2
9	23	4	0	1	0	0	2
10	28	4	0	1	1	0	2
11	15	2	1	1	1	0	2
12	05	3	0	1	1	1	3
13	10	3	0	1	0	0	2
14	45	4	0	2	0	1	3
15	04	3	1	1	1	0	2
16	50	3	1	2	1	0	2
17	67	3	1	2	1	0	2
18	24	4	0	1	1	1	4
19	08	4	0	1	0	1	3
20	12	3	0	1	1	0	2
21	05	4	0	1	0	1	3
22	07	4	1	1	1	0	2
23	08	4	0	2	0	1	4
24	08	4	1	1	0	0	2
25	20	3	0	1	1	1	3
26	10	3	0	1	1	0	2
27	33	3	1	1	1	0	2
28	30	4	1	3	0	0	2
29	39	3	0	2	1	1	2
30	35	4	0	2	1	0	2
31	15	3	0	3	0	0	2

1 = number of item
2 = q-value
3 = amount of tones in one subgroup
4 = misleading accent (1 = yes, 0 = no)
5 = amount of bits (see text)
6 = "no"-item = 1, "yes"-item = 0
7 = structure-forming variable (1 = pitch, 0 = length or intensity)
8 = amount of different tones

Appendix 6. Primary matrix for explaining the difficulty of the items. Version D. q-values obtained from 155 public school fourth-graders.

1	2	3	4	5	6	7	8
1	10	2	0	1	2	1	2
2	23	2	0	1	3	0	2
3	49	2	0	1	3	0	2
4	22	3	0	1	2	1	3
5	25	4	0	1	1	0	2
6	17	2	0	1	3	1	2
7	46	3	0	1	2	0	2
8	23	4	0	1	1	1	2
9	41	2	0	1	4	0	2
10	59	3	0	1	3	0	2
11	67	3	0	1	2	0	2
12	27	3	0	1	2	1	2
13	23	3	0	1	3	1	3
14	55	6	0	3	1	0	2
15	66	3	1	1	2	0	2
16	35	2	1	1	3	0	2
17	21	2	0	1	4	1	2
18	70	3	0	2	2	0	2
19	57	3	0	2	3	1	2
20	65	6	1	1	1	0	2
21	22	2	1	1	4	0	2
22	26	3	0	1	3	0	2
23	37	3	0	1	2	0	2
24	66	7	0	2	1	1	2
25	71	3	1	2	2	0	2
26	61	3	1	2	3	0	2
27	43	3	0	1	3	1	3
28	83	6	0	3	1	0	2
29	35	4	1	1	2	0	2
30	43	3	0	2	4	1	2
31	43	3	0	1	3	0	2
32	59	5	0	3	2	1	2
33	23	4	0	1	3	0	2
34	86	6	0	3	1	0	2
35	43	5	0	3	3	1	3
36	61	5	0	3	2	0	2

1 = number of item
2 = q-value
3 = amount of tones in one subgroup
4 = misleading accent (1 = yes, 2 = no)
5 = amount of bits (see text)
6 = amount of subgroups
7 = structure-forming variable (1 = pitch, 0 = length or intensity)
8 = amount of different tones

Appendix 7. Primary matrix for explaining the difficulty of the items. Version E. q-values obtained from 269 applicants for music training.

1	2	3	4	5	6	7	8
1	03	2	0	1	2	1	2
2	01	2	0	1	3	1	2
3	04	2	0	1	3	0	2
4	02	3	0	1	2	1	3
5	06	4	0	1	1	0	2
6	02	2	0	1	4	1	2
7	03	4	0	1	1	1	2
8	08	4	0	1	3	0	2
9	22	3	0	1	3	0	2
10	02	3	0	1	2	0	2
11	39	2	0	1	4	0	2
12	06	3	0	1	2	1	2
13	04	3	0	1	3	1	3
14	11	2	1	1	3	0	2
15	08	3	0	1	2	0	2
16	04	2	1	1	4	0	2
17	07	4	1	1	2	0	2
18	25	3	0	2	4	1	2
19	25	3	0	1	3	0	2
20	16	2	0	1	3	0	2
21	04	3	0	1	3	0	2
22	09	3	0	1	3	1	3
23	23	5	0	3	3	1	3
24	12	6	0	3	1	0	2
25	40	3	0	2	2	0	2
26	14	6	1	1	1	0	2
27	22	7	0	2	1	1	2
28	38	3	1	2	3	0	2
29	12	5	0	3	2	0	2
30	39	6	0	3	1	0	2
31	12	3	1	1	2	0	2
32	63	3	1	2	2	0	2
33	41	3	0	2	3	1	2
34	38	3	0	1	2	0	2
35	30	6	0	3	1	0	2

1 = number of item
2 = q-value
3 = amount of tones in one subgroup
4 = misleading accent (1 = yes, 0 = no)
5 = amount of bits (see text)
6 = amount of subgroups
7 = structure-forming variable (1 = pitch, 0 = length or intensity
8 = amount of different tones

Appendix 8. Primary matrix for explaining the difficulty of the items. Version G. q-values obtained from 79 public school fourth graders.

1	2	3	4	5	6	7
1	05	3	0	1	1	3
2	08	4	0	1	0	2
3	42	5	0	1	0	2
4	06	3	0	1	1	2
5	11	4	0	1	0	2
6	38	6	0	1	0	2
7	32	6	0	1	0	2
8	09	4	0	1	1	2
9	05	3	0	1	0	2
10	10	4	0	1	0	2
11	28	4	0	1	1	2
12	59	5	0	1	0	2
13	06	4	0	1	1	2
14	41	4	1	1	0	2
15	11	3	0	1	1	3
16	33	3	1	1	0	2
17	09	5	1	1	0	2
18	66	5	1	1	0	2
19	13	5	0	1	1	5
20	56	4	1	3	0	2
21	78	5	1	4	0	2
22	90	3	1	2	0	2
23	29	4	1	3	1	3
24	78	4	0	2	1	2
25	72	3	1	2	0	2
26	23	5	0	2	0	2
27	25	4	0	1	1	4
28	92	6	0	4	0	2
29	80	4	1	2	0	2
30	18	4	0	2	1	3
31	73	4	1	3	0	2
32	41	5	0	2	1	3

1 = number of item
2 = q-value
3 = amount of tones in one subgroup
4 = misleading accent
5 = amount of bits (see text)
6 = structure-forming variable (1 = pitch, 0 = length or intensity
7 = amount of different tones

Appendix 9. Correlation matrices for explaining the difficulty of test items.

Version A

	1	2	3	4	5	6
1 difficulty						
2 subgroup length	-.05					
3 misleading accent	.10	-.18				
4 amount of bits	.38	.24	-.27			
5 "no" item	.03	-.08	-.07	-.30		
6 pitch item	-.49	.19	-.34	-.02	.21	
7 amount of different tones	-.24	.36	-.25	-.15	.04	.56

Version C

	1	2	3	4	5	6
1 difficulty						
2 subgroup length	.08					
3 misleading accent	.33	-.10				
4 amount of bits	.54	.12	.18			
5 "no" item	.27	-.30	.24	-.16		
6 pitch item	-.15	.11	-.44	-.06	-.14	
7 amount of different tones	-.13	.39	-.32	-.05	-.20	.74

Appendix 9 continued

Version D

	1	2	3	4	5	6
1 difficulty						
2 subgroup length	.56					
3 misleading accent	.15	-.08				
4 amount of bits	.62	.67	-.14			
5 amount of subgroups	-.41	-.71	.02	-.32		
6 pitch item	-.36	-.02	-.37	.04	.12	
7 amount of different tones	-.21	-.00	-.17	-.00	.14	.47

Version E

	1	2	3	4	5	6
1 difficulty						
2 subgroup length	.15					
3 misleading accent	.13	-.06				
4 amount of bits	.49	.65	-.12			
5 amount of subgroups	-.00	-.71	.02	-.31		
6 pitch item	-.22	-.07	-.36	-.04	.14	
7 amount of different tones	-.18	.01	-.18	.02	.13	.50

Version G

	1	2	3	4	5
1 difficulty					
2 subgroup length	.22				
3 misleading accent	.51	-.18			
4 amount of bits	.65	.19	.40		
5 pitch item	-.41	-.26	-.43	-.15	
6 amount of different tones	-.29	-.01	-.24	-.10	.59

Appendix 10. Factor matrix, version A of the acoustic structuring test. Three first factors of a principal axes solution.

item	I	II	III	h^2
1	.268	-.164	-.186	.133
2	.334	-.217	.114	.172
3	.250	-.234	.230	.171
4	.313	-.088	.033	.107
5	.303	.043	.110	.106
6	.391	-.343	.123	.286
7	.242	.140	.063	.082
8	-.120	-.170	.212	.088
9	.201	-.129	.206	.100
11	.218	.308	-.199	.182
12	.339	.216	.252	.225
14	.118	.036	-.049	.018
15	.063	-.048	.054	.009
16	.157	-.161	.006	.051
17	.169	-.156	-.219	.101
18	.357	-.150	-.117	.164
19	.133	-.154	-.058	.045
20	.232	.269	.037	.127
22	.267	.357	-.124	.214
23	.212	-.243	-.038	.105
24	.374	-.243	.017	.199
25	.074	-.229	-.112	.071
26	.196	.024	.105	.050
27	.086	-.240	-.266	.136
28	.294	.057	-.152	.113
29	.273	.020	-.019	.075
30	.293	.280	.300	.254
31	.236	.285	-.114	.150
32	.045	.006	-.233	.056
33	.300	.282	-.226	.221
34	.135	.022	-.196	.057
35	.253	-.238	-.038	.122
36	.310	.043	-.305	.190
37	.396	-.036	.118	.172
38	.320	.353	.169	.256
39	.266	-.088	-.009	.079
40	.139	.065	.093	.032
eigen-values	2.366	1.431	0.923	4.720

Appendix 11. Factor matrix, version C of the acoustic structuring test. Three first factors of a principal axes solution.

item	I	II	III	h^2
1	.289	.162	-.002	.110
2	.108	.230	-.162	.091
3	.251	-.255	-.187	.168
4	.217	.126	.039	.065
5	.233	-.086	.118	.076
6	.267	.330	-.056	.183
7	.371	.002	-.218	.185
8	.046	.319	-.112	.116
9	-.037	.102	.033	.013
10	.469	-.053	-.171	.252
11	.270	-.261	.117	.155
12	.322	-.241	-.130	.179
13	.086	-.018	-.126	.024
14	.031	.139	.359	.149
15	.172	-.150	.229	.105
16	.209	-.075	.252	.113
17	.026	-.241	-.273	.133
18	.287	-.008	.017	.083
19	.223	.331	.057	.162
20	.495	-.016	-.131	.263
21	.291	.284	.022	.165
22	.377	-.092	.107	.162
23	.089	.285	-.202	.130
24	.085	.137	.184	.060
25	.263	-.092	-.064	.082
26	.308	-.169	.305	.216
27	.191	-.165	.076	.069
28	.077	.197	.244	.104
29	.380	.091	.057	.156
30	.193	-.175	.045	.070
31	.207	.239	.000	.100
eigen-values	2.006	1.121	0.812	3.939

Appendix 12. Factor matrix, version D of the acoustic structuring test. Three first factors of a principal axes solution.

item	I	II	III	h^2
1	.353	-.138	.113	.157
2	.546	-.319	.057	.403
3	.103	-.091	.490	.259
4	.657	.053	-.091	.443
5	.597	.051	-.313	.457
6	.628	-.289	.012	.478
7	.274	.185	-.249	.171
8	.611	-.082	-.128	.397
9	.349	-.276	.345	.317
10	.256	.282	-.011	.145
11	.057	.093	-.144	.032
12	.399	.381	.031	.305
13	.286	.442	.168	.306
14	.489	.040	-.278	.318
15	-.029	.442	-.071	.201
16	.105	.068	.080	.022
17	.641	-.225	.156	.486
18	-.132	.427	.144	.220
19	-.144	.146	.149	.064
20	.472	-.054	-.315	.325
21	.496	-.340	.152	.385
22	.429	-.002	-.014	.184
23	.335	.375	.009	.253
24	.406	.007	-.258	.231
25	.208	.160	.067	.073
26	.203	.263	-.079	.117
27	.146	.166	.302	.140
28	.011	-.000	-.220	.048
29	.434	.251	.120	.266
30	.111	.008	.324	.117
31	.199	.144	.113	.073
32	.092	.136	.077	.033
33	.546	-.172	.081	.335
34	.047	-.010	-.201	.043
35	.347	.312	.182	.251
36	.430	.203	.201	.266
eigen-values	5.088	1.872	1.364	8.324

Appendix 13. Factor matrix, version E of the acoustic structuring test. Three first factors of a principal axes solution.

item	I	II	III	h^2
1	.180	.044	-.116	.063
2	.270	-.311	.179	.373
3	.248	-.113	.113	.096
4	.179	.020	.029	.129
5	.274	.317	-.015	.235
6	.276	-.024	-.073	.107
7	.500	-.227	.150	.503
8	.323	.064	.195	.177
9	.316	.103	-.274	.219
10	.363	-.251	.156	.277
11	.328	.065	-.473	.339
12	.376	-.127	.066	.165
13	.370	-.440	.105	.372
14	.101	.079	-.138	.042
15	.362	-.151	-.036	.155
16	.197	-.104	.064	.130
17	.205	-.161	-.098	.089
18	.118	-.011	-.261	.086
19	.238	-.052	-.335	.210
20	.256	.207	-.496	.368
21	.216	-.048	-.065	.066
22	.301	-.270	.147	.216
23	.441	.106	.112	.220
24	.679	.225	.087	.528
25	.250	-.342	-.033	.181
26	.578	.227	.174	.417
27	.481	.249	.057	.297
28	.007	-.365	-.010	.159
29	.403	.148	.011	.189
30	.499	.275	.066	.330
31	.404	-.107	-.061	.222
32	-.159	-.366	-.118	.177
33	.330	-.312	-.211	.251
34	.159	-.293	-.345	.232
35	.402	.220	.133	.229
eigen-values	3.997	1.642	1.208	7.848

Appendix 14./Factor matrix, version G of the acoustic structuring test. Three first factors of a principal axes solution.

item	I	II	III	h^2
1	.530	-.527	.084	.566
2	.269	.243	.092	.140
3	.256	.252	.434	.318
4	.563	-.004	-.172	.346
5	.355	-.273	-.406	.365
6	.396	-.106	.277	.245
7	.562	-.030	-.027	.317
8	.295	-.027	.125	.103
9	.530	.298	-.353	.495
10	.505	.397	-.189	.448
11	.568	.053	.186	.361
12	.383	.122	.219	.209
13	.428	-.175	.064	.218
14	-.147	.051	-.320	.127
15	.142	.341	-.221	.185
16	.259	-.037	-.197	.107
17	.251	-.383	-.188	.246
18	.038	-.282	.063	.085
19	.583	-.395	.163	.523
20	.218	.113	-.026	.061
21	-.232	.335	-.175	.197
22	-.042	.055	.336	.117
23	.217	-.071	.240	.110
24	.099	-.046	-.132	.029
25	.262	.225	-.022	.120
26	-.104	-.123	-.126	.042
27	.479	.320	.012	.333
28	-.075	.178	.193	.075
29	.221	-.219	-.189	.133
30	.503	.223	-.180	.335
31	.058	.083	.103	.021
32	.185	.306	.254	.192
eigen-values	3.938	1.818	1.411	7.167

Research Bulletin, Institute of Education, University of Helsinki

xNo 1 The Effectiveness of Punishments and the School Climate by OLLI SIPINEN. May, 1957. 10 pp.

No. 2 The Balance of Instructional Situations by MATTI KOSKENNIEMI. Nov., 1957. (reprint 1972). 7 pp.

xNo. 3 On Deceit Practiced in Finnish Grammar Schools at Written Tasks by OLAVI PÄIVÄNSALO. Nov., 1957. 5 pp.

No. 4 Comparison between Reading Comprehension and Reading Aloud by KAISA HÄLINEN. March, 1958. 9 pp.

xNo. 5 Ten Test-Papers in the Swedish Language as Graded by Ten Teachers by VÄINÖ HEIKKINEN & ANITA HAVELA. Sept., 1958. 5 pp.

xNo. 6 The Awarding of School Marks by TOIVO VAHERVUO. Sept., 1958. 5 pp.

xNo. 7 Some Comments on the Uniformity of the 'Experimental Period' in the History of Soviet Education by MATTI KOSKENNIEMI. Feb., 1959. 18 pp.

xNo. 8 Motives for the Choice of the Teaching Profession Correlated with Success in the Entrance Examination and in Student Teaching by MATTI KOSKENNIEMI. Feb., 1960. 11 pp.

xNo. 9 One's Own Teacher and the Choice of Teaching Profession by MATTI KOSKENNIEMI & EIVI LIHTONEN. Sept., 1960. 14 pp.

xNo. 10 The Curriculum of School Mathematics and Mathematical Performance. An Experimental Study by PAAVO MALINEN. Sept., 1961. 50 pp.

No. 11 Prognostic Value of Some Psychological Test Results Correlated with Student Teaching Marks by MARJA-LEENA KÄYHTY. Dec., 1961. 10 pp.

xNo. 12 Factors Constituting Classroom Climate by OLLI SIPINEN. March, 1963. 10 pp.

xNo. 13 On the Changes in the Educational Attitudes of Young Teacher by VÄINÖ HEIKKINEN. March, 1963. 14 pp.

xNo. 14 Teacher Training in Retrospection. A Preliminary Report by VÄINÖ HEIKKINEN. Feb., 1964. 21 pp.

No. 15 Placement of Young Elementary School Teachers by MATTI KOSKENNIEMI. May, 1965. (reprint 1972). 9 pp.

xNo. 16 Some Problems in Combining the Individual Tutors' Marks into a Final Mark in Teaching Ability by JOHANNES ALIKOSKI. March, 1966. 25 pp.

No. 17 The Finnish Senior Secondary Research Project I. General Presentation of the Project and Its Original Subjects by ANNA-LIISA SYSIHARJU. Sept., 1967. 31 pp.

xNo. 18 The Finnish Senior Secondary Research Project II. Departure from Junior Secondary and Transfer to Senior Secondary by ANNA-LIISA SYSIHARJU. Oct., 1967. 52 pp.

No. 19 Relationships between Absences from School and Certain Factors Related to School. A Study on Elementary School Grades Three to Six by PERTTI KANSANEN. Nov., 1967. 10 pp.

No. 20 On the Anxiety Associated with School Attendance and the Grammar School Entrance Examination by EEVA PATJAS. May, 1968. 16 pp.

No. 21 School Achievement and Personality. I: Design and Hypotheses, IV: Results and Discussion by ERKKI A. NISKANEN. Oct., 1968. 51 pp.

No. 22 School Achievement and Personality. II: Operations at the Variable Level by ERKKI A. NISKANEN. Oct., 1968. 124 pp.

No. 23 School Achievement and Personality. III: Operations at the Factor Level by ERKKI A. NISKANEN. Oct., 1968. 39 pp.

No. 24 The Effectiveness of Two Methods of Teaching English as a Foreign Language in Some Finnish Secondary Schools by DANIEL J. CASEY. Nov., 1968. 37 pp.

No. 25 The Learning of Elementary Algebra. An Empirical Investigation of the Results of Learning in a Simplified School Learning System by PAAVO MALINEN. Jan., 1969. 85 pp.

No. 26 Investigations into the Instructional Process. I. Some Methodological Problems by MATTI KOSKENNIEMI and ERKKI KOMULAINEN in collaboration with Anna-Kaarina Falck and Pentti Holopainen. Oct., 1969. (reprint 1972). 35 pp.

No. 27 Investigations into the Instructional Process. II. Objectivity of Coding in a Modified Flanders Interaction Analysis by ERKKI KOMULAINEN. Dec., 1970. 26 pp.

No. 28 Investigations into the Instructional Process. III. P-technique Treatment of Observational Data by ERKKI KOMULAINEN. Jan., 1971. 41 pp.

No. 29 Investigations into the Instructional Process. IV. Teaching as a Stochastic Process by ERKKI KOMULAINEN. Feb., 1971. 24 pp.

No. 30 Investigations into the Instructional Process. V. Experiences with the Bellack Classification System by KAI KARMA. Jan., 1972. 16 pp.

x No. 31 Investigations into the Instructional Process. VI. Group Work as a Social Interaction Process. A Case Approach by PENTTI HOLOPAINEN. May, 1972. 34 pp.

No. 32 Investigations into the Instructional Process. VII. Non-verbal Behaviour in Group Work. Observational Technique and Typology of Pupils' Work by MARJA MARTIKAINEN. Sept., 1972. 18 pp.

x No. 33 English School Achievements and Some Student Characteristics. I. On the Relationships of Personality and Intelligence Variables to English School Achievements by ANNA-LIISA LEINO. Oct., 1972. 50 pp.

No. 34 Investigations into the Instructional Process. VIII. On the Problems of Variable Construction from Flanders' Interaction Matrix with Special Emphasis on the Stochastic Nature of Classroom Communication by ERKKI KOMULAINEN. Feb., 1973. 78 pp.

No. 35 Zur Problematik des Unterrichts by CHR. PETROW CHRISTOW. April, 1973. 11 pp.

No. 36 Investigations into the Instructional Process. IX. Pupils' Goal-related Behavior during the Instructional Interaction by MATTI KOSKENNIEMI & PENTTI HOLOPAINEN. June, 1973. 17 pp.

x No. 37 The Behaviour of Supervisors assessed by Student Teachers by SEPPO KONTIAINEN. Nov., 1973. 186 pp. Out of print.

No. 38 The Ability to Structure Acoustic Material as a Measure of Musical Aptitude. I. Background Theory and Pilot Studies by KAI KARMA. Nov., 1973. 19 pp.

x No. 39 Investigations into the Instructional Process. X. Report of the DPA Helsinki, Phase One by MATTI KOSKENNIEMI & ERKKI KOMULAINEN. Feb., 1974. 35 pp.

x No. 40 English School Achievements and Some Student Characteristics II. On the Relationships of Motivational and Auditory Variable to English School Achievements. Final Analyses by ANNA-LIISA LEINO. March, 1974. 48 pp.

No. 41 Foreign Language Curriculum. An Investigation into Objectives, Achievements and Student Characteristics by ANNA-LIISA LEINO. June, 1974. 27 pp.

No. 42 DPA Helsinki System for Describing Instructional Process. Manual. Prepared and edited on behalf of the Research Team DPA Helsinki by MATTI KOSKENNIEMI, KAI KARMA and MARJA MARTIKAINEN. Nov., 1974. 34 pp.

No. 43 The Ability to Structure Acoustic Material as a Measure of Musicai Aptitude. 2. Test Construction and Results by KAI KARMA. Feb., 1975. 21 pp.

No. 44 A Comparison of the Importance of Objectives in Foreign Language Instruction as Rated by In-Service Teachers, Student Teachers and Pupils by ANNA-LIISA LEINO. June, 1975. 35 pp.

No. 45 Implementation of Some Affective Objectives in Different Foreign Language Teaching Situations as Rated by Student Teachers by ANNA-LIISA LEINO. Nov., 1975. 45 pp.

No. 46 The Need Felt for In-Service Education by Foreign Language Teachers by ANNA-LIISA LEINO. March, 1976. 53 pp.

No. 47 The Ability to Structure Acoustic Material as a Measure of Musical Aptitude. 3. Theoretical Refinements by KAI KARMA. Aug., 1976. 32 pp.

No. 48 Structural Elements of the Instructional Process: Activity Forms and Content. DPA Helsinki Investigations I by MATTI KOSKENNIEMI. Dec., 1977. 36 pp.

No. 49 Research in Teaching. Papers presented at the symposion arranged by the Academy of Finland for evaluation of the DPA Helsinki Project, Helsinki, October 26th and 27th 1977. DPA Helsinki Investigations II edited by Erkki Komulainen and Matti Koskenniemi. June, 1978. 152 pp.

No. 50 An Introduction to the Content of Foreign Language Instruction by Anna-Liisa Leino. June, 1978. 80 pp.

No. 51 The Ability to Structure Acoustic Material as a Measure of Musical Aptitude. 4. Experiences with Modifications of the Acoustic Structuring Test by Kai Karma. Dec., 1978. 51 pp.

x) Out of print.